

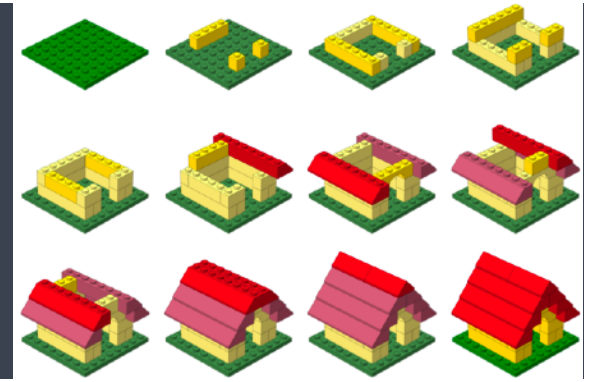
# Introduction to R

Harvard Chan Bioinformatics Core

<https://tinyurl.com/hbc-r-online>

Sponsored by DF/HCC, CFAR, and HMS Foundry

# Learning Objectives



- ✓ Comfortably use RStudio (a graphical interface for R)
- ✓ Fluently interact with R using RStudio
- ✓ Become familiar with R syntax
- ✓ Understand data structures in R
- ✓ Inspect and manipulate data structures
- ✓ Install packages and use functions in R
- ✓ Visualize data using *ggplot2*
- ✓ Utilize pipes, tibbles and functions from the Tidyverse package suite

# Exit survey

<https://tinyurl.com/r-workshop-hbc>

# Interested in additional training?

<https://bioinformatics.sph.harvard.edu/current-bioinformatics-topics-workshops>

---

## Short workshops: Current Topics in Bioinformatics

These workshops are free and open to all researchers at Harvard University and affiliated institutions.

- **Workshops** on bioinformatics methods & related skills.
- Once a month for 3 hours
- Hands-on workshops - be prepared with your MAC or Windows computer
- **Free and open to everyone at Harvard University and its affiliates**
- Will typically meet the **third Wednesday of the month from 1-4 online via Zoom** (meeting time is subject to change-please check)

# Big data? Big computer!

## The skill set you need to succeed

<https://hbctraining.github.io/Training-modules/>

---

**Spring 2024 schedule - Big data? Big computer! The skill set you need to succeed:**

Topic	Pre-requisites	Date/Time	Time	Registration
Shell Tips and Tricks on O2	<a href="#">Basic Shell</a>	5/15/24	1 – 4pm	<a href="#">Register!</a>

# Harvard Catalyst Online Resource

<https://projects.iq.harvard.edu/hcatresource>



HARVARD UNIVERSITY

HARVARD.EDU

Harvard Catalyst Introduction to R:  
*An online, hands-on training resource for learning the basics of R*

[Contact](#)



HARVARD  
CATALYST

Harvard Clinical & Translational Science Center

HOME

Lessons

Faculty

Supplemental Resources

## Welcome to Introduction to R

This **online, hands-on learning resource** will introduce you to using R and RStudio. R is a simple programming environment that enables the effective handling of data, while providing excellent graphical support. RStudio is a tool that provides a user-friendly environment for working with R. This resource is intended to provide both basic R programming knowledge and information on utilizing R to increase efficiency in data analysis.

This comprehensive online learning resource was created in collaboration between [Harvard Catalyst](#) and the [Harvard Chan Bioinformatics Core](#). It includes a series of videos explaining fundamental concepts in R and demonstrates the application through live coding. It is geared toward those interested in learning the basics of R for reproducible data wrangling and visualizations (ggplot2), and/or performing data analyses that require a basic knowledge of R.

Resource lessons address the following:

- **R syntax:** Understanding the different 'parts of speech' in R, and introducing variables and functions, demonstrating how functions work, and modifying arguments for specific use cases.
- **Data structures in R:** Explaining the classes of data structures and the types of data used by R.
- **Data inspection and wrangling:** Reading in data from files, and using indices and various functions to subset and create datasets (including the tidyverse suite of packages).
- **Visualizing data:** Visualizing data using plotting functions from the external package ggplot2.
- **Exporting data and graphics:** Generating new data tables and plots for use outside of the R



# Data Management

- HMS Data management -

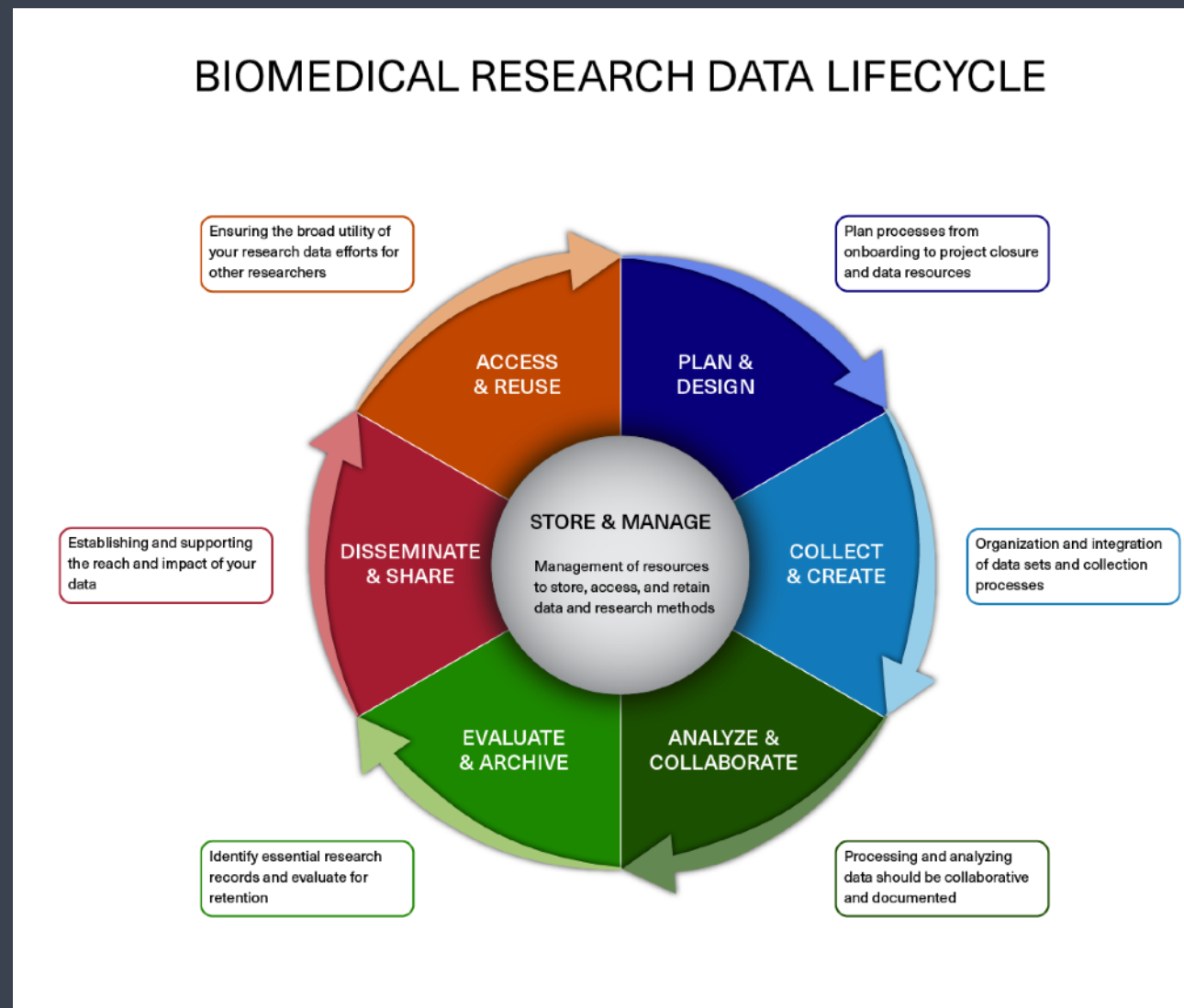
Webpage: <https://datamanagement.hms.harvard.edu/>

[Click here to sign up for data management related emails](#)

- Harvard-wide Research Data Management -

<https://researchdatamanagement.harvard.edu/>

# Data Management Short Workshops



<https://datamanagement.hms.harvard.edu/about/news-events/rdmwg-calendar>









# Data Management Short Workshops

## Spring 2024 Data Lifecycle Training






### Plan & Design

- April 17**  Writing a Data Management and Sharing Plan for Grant Applications
- April 24**  Principles of Data Stewardship
- May 1**  Don't Leave Yet! Research Data Offboarding Done Right
- May 8**  Don't Leave Yet! Research Data Offboarding Done Right




### Collect & Analyze



- February 21**  Basic Shell
- March 6**  Using Metadata to Find, Interpret & Share Your Data
- March 13**  Intro to MATLAB
- March 13**  Introduction to GIS
- April 17**  Intermediate Shell
- May 15**  Active Project Version Control with Git

### Store & Evaluate

- February 7**  The When, Where, and How of Data Storage
- February 27**  Introduction to the General Records Schedule
- March 20**  Managing Data Transfer Between Collaborators
- March 26**  Managing Your Paper Records
- May 7**  Managing Your Electronic Records

### Share & Publish

- January 16**  Data Sharing and Reuse on the Vivli Platform
- February 13**  Introduction to Harvard Dataverse
- February 15**  Research Management Using the Open Science Framework
- April 10**  Achieving FAIR Data: Selecting a Repository for Your Data

-  In-person
-  Virtual



**Learn More & Register:** [bit.ly/rdmwg-calendar](https://bit.ly/rdmwg-calendar)



<https://datamanagement.hms.harvard.edu/about/news-events/rdmwg-calendar>

# Bioinformatics Community Networking Breakfast!



- *Free and open to the Harvard community*
- *Food is first-come-first-served*
- *No computers!*
- *Quarterly* (Next meeting TBA)

More info: <http://bioinformatics.sph.harvard.edu/breakfast/>

# Get (stay) in touch with us!

*Sign up for our mailing list:*

<https://tinyurl.com/hbc-training-mailing-list>

*Training email:* [hbctraining@hsph.harvard.edu](mailto:hbctraining@hsph.harvard.edu)

*Consulting email:* [bioinformatics@hsph.harvard.edu](mailto:bioinformatics@hsph.harvard.edu)

**Twitter:** [@bioinfocore](https://twitter.com/bioinfocore)